Construction of high-resolution physical maps for large plant genomes

S.F. Kianian[1], O. Riera-Lizarazu[2], Y.Q. Gu[3,4], M-C. Luo[4], A.M. Denton[1], and G.R. Lazo[3]

[1]North Dakota State University,[2]Oregon State University, [3] USDA-ARS Albany CA & [4]University of California-Davis

DBI-0822100

Long-term goal and specific objectives

We have devised a simple and an elegant method of developing high-resolution RH physical maps for the 4.2 Gb D-genome of hexaploid wheat. This transformative technology allows the construction of a fine marker scaffold that bridges the resolution gap between meiotic and BAC-based physical mapping. These maps will be valuable in the analysis of biologically and agronomically important genes, comparison with the rice, maize and other grass genomes, and an unprecedented look at gene organization and chromosome evolution of these important crop plant genomes. Once optimized, this approach can easily be applied to most plant species, including many important polyploid crops or those that lack substantial genomics resources, thereby changing the paradigm for the construction of physical maps, complete genome sequencing, and comparative analyses. The specific objectives of this project are:

A. Develop radiation hybrid mapping populations for D-genome chromosomes of Chinese Spring wheat (reference hexaploid wheat) and *Aegilops tauschii* Accession AL8/78 (wheat species with the most advanced physical mapping information)

i. Develop populations of RH lines for each of the two D-genomes at two different radiation levels
ii. Characterize the RH mapping populations with respect to marker retention and breakage frequency and adjust panel size for optimum mapping resolution

B. Develop RH maps of varying resolution for D-genome chromosomes and align BAC contigs to the maps using automated high-throughput protocols

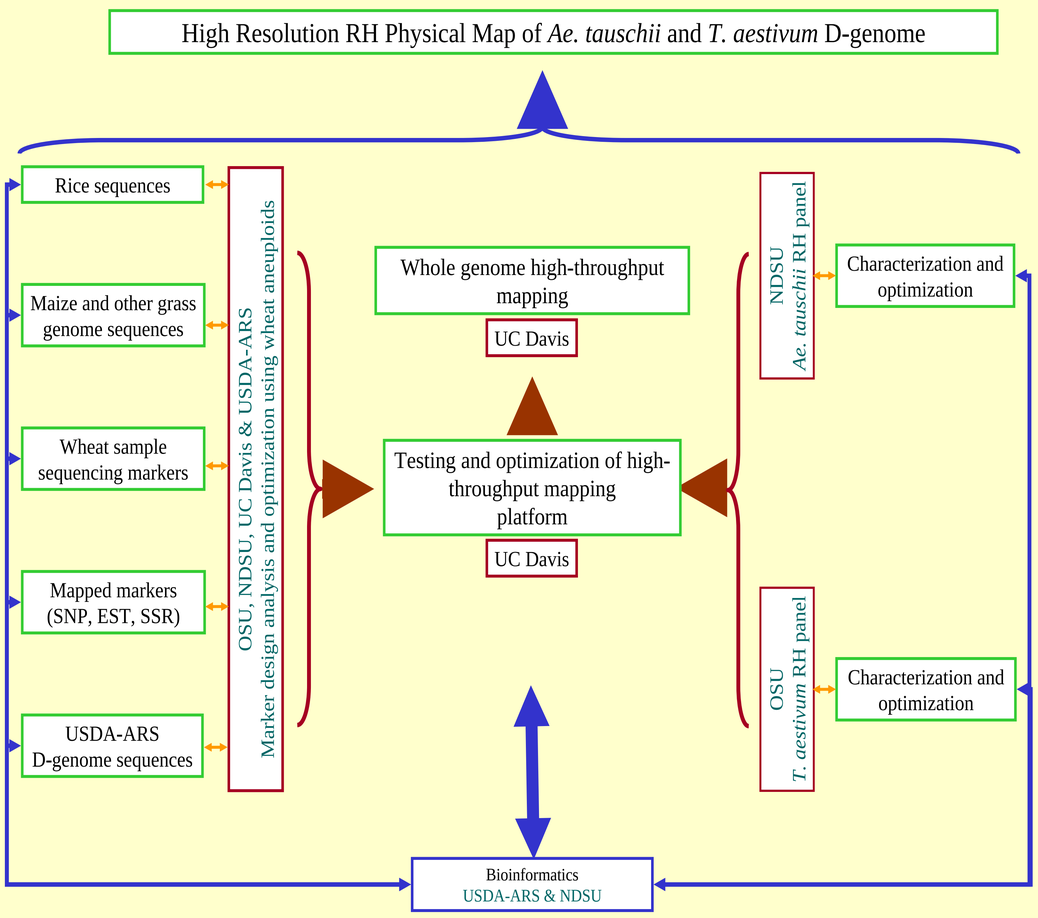
i. Develop ~10,000 markers derived from various sources (majority retro-element junctions) for mapping
ii. Optimize a cost effective and high-throughput approach for mapping the necessary markers
iii. Develop maps of D-genome chromosomes populated with ~8,000 loci using high- and mid-resolution RH[1] panels as well as wheat aneuploid stocks (low resolution controls) by high-throughput genotyping protocols
iv. Simultaneously, align *Ae. tasuchii* BAC clones and contigs, from libraries estimated at 5X genome coverage, utilizing a high-throughput approach
v. Determine gene distribution along the chromosomes, and construct detailed comparative maps between the two D-genomes and with barley, rice, maize and other model grass species

C. Develop the database and bio-informatics tools for efficient access to and utilization of the resources generated in this project

i. Timely and effective dissemination of experimental data
ii. Tools to assist in the experimental process

A. Integrate the knowledge and resources of this project into teaching/training of students at all levels

i. Integrate project methods into existing high school and undergraduate training programs
ii. Train a new generation of scientists at the graduate and postdoctoral level



The overall project organization

This schema represents project organization according to the major activities**.** All encompassing activities relate to education, outreach and bioinformatics. These three components will permeate all work being conducted by this group.

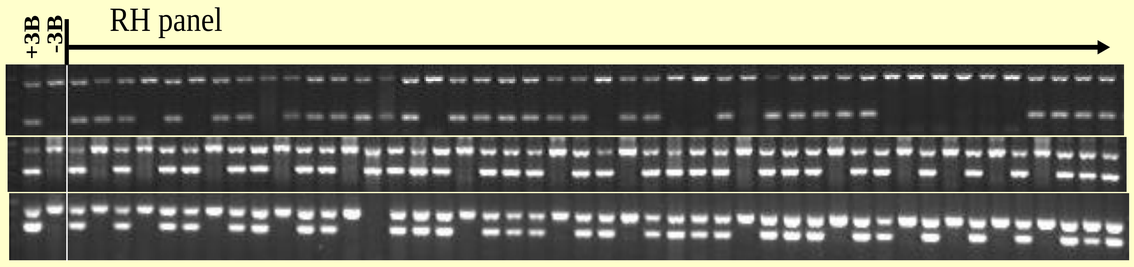


Figure 1: Three REJ markers (top cfp1869, middle cfp1444 and bottom cfp1589) mapped on chromosome 3B RH panel. In all PCR reactions a conserved gene primer set for chromosome 1A (upper band) is multiplexed with a REJ primer set (lower band) for chromosome 3B. The first lane is the chromosome 3B parental control and second lane is the line missing chromosome 3B.

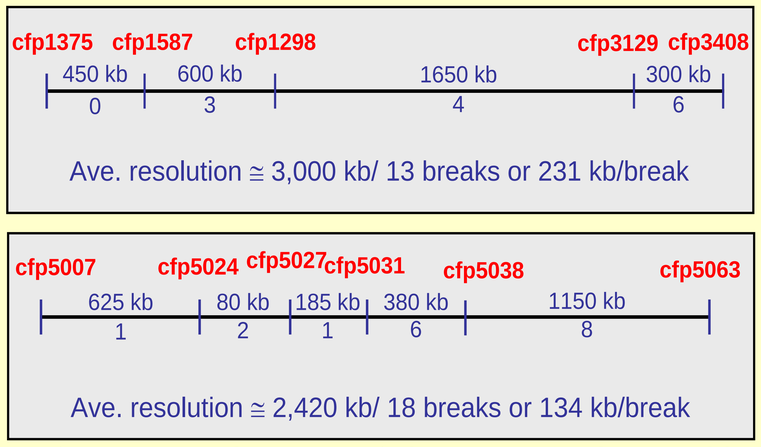


Figure 2: Resolution of two sequenced segments of chromosome 3B determined by RH mapping. REJ markers mapped on the panel are identified as cfpxxxx, distances between loci are presented in Kb above the map and number of breaks per given segment below. Average resolution is calculated by dividing the total length of a given segment by the total number of breaks identified. Overall average resolution calculated using data from all three segments is ~200Kb

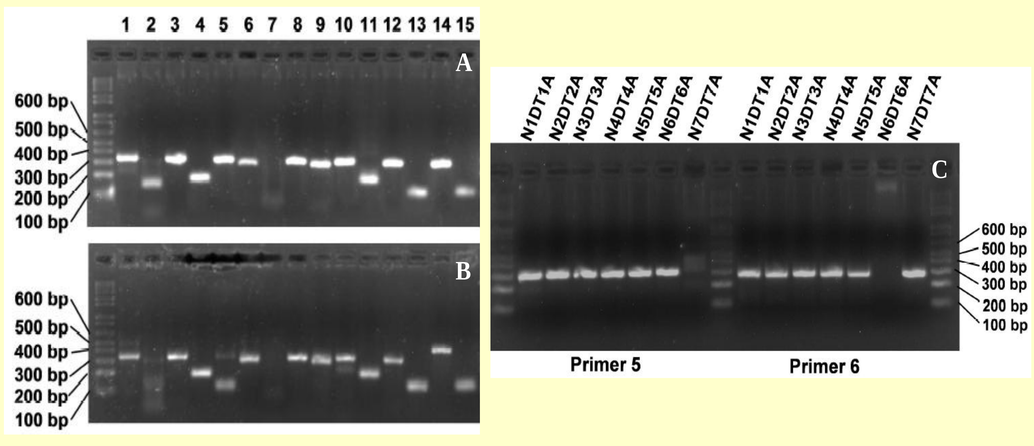


Figure 3: Amplification products resulting from primers designed from *Ae. tauschii* BES REJs; A) products in *Ae. tauschii*, B) products in *T. aestivum* CS, and C) products in CS aneuploid stocks nullisomics for a D-genome chromosome (e.g. N1D) and tetrasomic for the corresponding homoeologous chromosome in the A-genome (e.g. T1A). Based on the lack of amplification in CS aneuploids primer 5 is located on chromosome 7D and primer 6 on chromosome 6D.

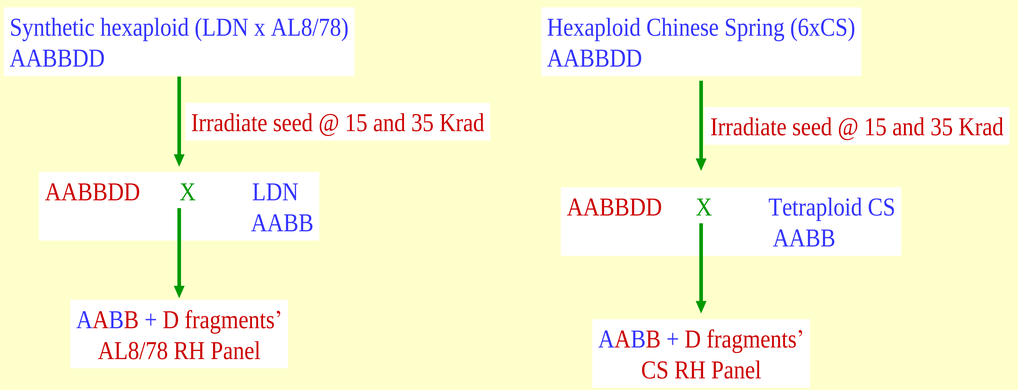


Figure 4: Scheme for the production of the D-genome RH mapping populations. Two approaches are identical except for the D-genome source and the tetraploid (AABB) recipient parent. Note that the D-genome in the case of the synthetic is the same genome used for the production of the *Ae. tauschii* physical map. The AABB genome backgrounds are nearly isogenic in both cases.